N89-19850

AN INTELLIGENT MULTI-MEDIA HUMAN-COMPUTER DIALOGUE SYSTEM

J.G. Neal, K.E. Bettinger, J.S. Byoun, Z. Dobes, & C.Y. Thielman
Calspan-UB Research Center (CUBRC)
4455 Genesee Street
Buffalo, New York 14225

ABSTRACT

Sophisticated computer systems are being developed to assist in the human decision-making process for very complex tasks performed under stressful conditions. The human-computer interface is a critical factor in these systems. The human-computer interface should be simple and natural to use, require a minimal learning period, assist the user in accomplishing his task(s) with a minimum of distraction, present output in a form that best conveys information to the user, and reduce cognitive load for the user. In pursuit of this ideal, the Intelligent Multi-Media Interfaces project, sponsored by the Defense Advanced Research Projects Agency and monitored by the Rome Air Development Center, is devoted to the development of interface technology that integrates speech, natural language text, graphics, and pointing gestures for human-computer dialogues. The objective of the project is to develop interface technology that uses the media/modalities intelligently in a flexible, context-sensitive, and highly integrated manner modelled after the manner in which humans converse in simultaneous coordinated multiple modalities. As part of the project, a knowledge-based interface system, called CUBRICON (CUBRC Intelligent CONversationalist) is being developed as a research prototype. The application domain being used to drive the research is that of military tactical air control.

1. INTRODUCTION

As the number and sophistication of military information processing systems rapidly increases, the impact on human operational users must be considered very carefully. Typically, large amounts of information must be communicated for use by the human operator in performing time-critical decision-making tasks for command and control functions. The problem is to make such sophisticated systems easy for military operators to use quickly and efficiently. Modern information processing and decision-aiding systems require a full range of communication media to facilitate interaction and provide the increased bandwidth for information transfer with the human user. The human-computer interface must not only take advantage of multiple media/modalities, but must make use of the synergistic properties of these media to minimize the user's cognitive workload: (1) the human-computer interface system must provide the user with input media/modalities that are natural and efficient for the user; and (2) the use of multiple media/modalities must be applied to the problem of presenting output information to the user in a manner that maximizes user comprehension; the manner of presentation must be based on knowledge of the application domain, the characteristics of the information, the discourse context, the user's task, and respected human factors guidelines.

Presently, there is no computer system that can meet the above requirements. Knowledge-based understanding and generation of information by a computer system in multiple media/modalities has recently begun to be investigated [Neches & Kaczmarek, 1986; Kobsa et al., 1986; Arens et al., 1988; Neal & Shapiro, 1988; Neal et al., 1988; Sullivan & Tyler, 1988]. The research discussed in this paper is part of the Intelligent Multi-Media Interfaces (IMMI) project [Neal & Shapiro, 1988; Neal et al., 1988] which is dedicated to the development of intelligent multi-media interface technology that integrates speech, natural language text, graphics, and pointing gestures for human-computer dialogues. The objective of the project is to develop interface technology that uses the media/modalities intelligently in a flexible, context-sensitive, and highly integrated manner modelled after the way in which humans converse in simultaneous coordinated multiple modalities. As part of the project, a knowledge-based interface system, called CUBRICON (the CUBRC Intelligent CONversationalist), is being developed as a proof-of-concept prototype. Although the IMMI project is a basic research project, an application task domain of military tactical air control is being used to drive the research.

This paper discusses the CUBRICON human-computer dialogue system and the multi-media communication methodology which forms the foundation of the system. Section 2 presents a brief overview of the CUBRICON system. Section 3 discusses the knowledge sources used by the system. Section 4 focuses on multi-media input processing while Section 5 focuses on the generation of multi-media output. Subsequent sections outline the future direction for the CUBRICON project, summarize the main ideas of this paper, provide acknowledgements, and list references.

2. OVERVIEW

The model upon which the CUBRICON system is based is that of two people talking in front of a blackboard or other graphics display device(s). People in such a situation use various combinations of modalities for very effective communication. The modali-

This research is supported by the Defense Advanced Research Projects Agency and monitored by the Rome Air Development Center under Contract No. F30603-87-C-0136.

ties include spoken natural language, written text, pointing gestures, drawings, and tables. Thus, CUBRICON is intended to imitate, to a certain extent, the ability of humans to simultaneously accept input from different sensory devices (such as eyes and ears), and to simultaneously produce output in different media (such as voice, pointing motions, and drawings).

CUBRICON accepts input from three input devices: speech input device, keyboard, and mouse device pointing to objects on a graphics display. Output is produced for three output devices: color-graphics display, monochrome display, and speech output device. The CUBRICON software is implemented on a Symbolics Lisp Machine using the SNePS semantic network processing system [Shapiro, 1979; Shapiro & Rapaport, 1986], an ATN parser/generator [Shapiro, 1982] and Common Lisp. Speech recognition is handled by a Dragon Systems VoiceScribe 1000. Speech output is produced by a DECtalk speech production system.

Every intelligent entity requires a considerable amount of knowledge upon which to base its decision-making processes. Certain knowledge sources have been identified as essential to CUBRICON's multi-media communication ability. These knowledge sources are discussed in the next section.

3. KNOWLEDGE SOURCES

In order for CUBRICON to perform the critical functions of a human-computer interface, the following knowledge sources were found to be essential: (1) domain-specific and related interface knowledge, (2) multi-media language knowledge, (3) the discourse context, (4) the user/task model, and (5) human factors guidelines for enhancing human understanding, and (6) information characteristics.

3.1 Domain-Specific And Related Interface Knowledge

CUBRICON includes a knowledge base containing domain-specific and interface information. The domain-specific information is applicable to the mission planning task domain of the target application system. Task domain entities include airbases, surface-to-air missile (SAM) systems, fuel storage facilities, and targets. The knowledge base also includes essential information concerning how to present or express the various entities via the system's verbal/graphic language. This information includes the words and symbols used to express any given entity, which which symbols are appropriate under which conditions, and when particular colors are to be used.

3.2 Multi-Media Language Knowledge

CUBRICON's multi-media langauge is defined by the combination of its lexicon and grammar. A lexicon is the collection of all morphines, tokens, or signals that carry meaning in a given language. CUBRICON's lexicon consists of words, graphic figures, and pointing signals. The grammar defines how the morphemes, tokens, and signals of the lexicon can combine to form legal composite language structures. An example of a multi-media language structure that is legal according to the CUBRICON grammar is a noun phrase consisting of the typical linguistic syntax, accompanied by one or more pointing signals (pointing to objects on a graphics display).

3.3 The Discourse Model

The attentional discourse focus space [Grosz, 1978, 1986; Sidner, 1983; Grosz and Sidner, 1985] is a key knowledge structure that supports continuity and relevance in dialogue. The CUBRICON system tracks the attentional discourse focus space of the dialogue carried out in multi-media language and maintains a representation of the focus space in two structures: (1) a main focus list and (2) a set of ancillary focus lists called virtual displays. The main focus list includes those entities and propositions that have been explicitly expressed (by the user or by CUBRICON) via natural language, pointing, highlighting, or blinking. A virtual display is a list of all the objects that are "in focus" because they are visible in a given window on one of the displays. CUBRICON maintains one virtual display per window.

3.4 User/Task Model

A user/task model is essential as a basis for judging the relevance and importance of information items to the user. Carberry [1987] provides a brief summary of current research on user modeling. CUBRICON's user/task model includes (1) a task hierarchy and (2) an entity rating module.

The task hierarchy is a decomposition of the user's main tasks into subtasks. In a task oriented application system, there is usually some a priori knowledge of the task hierarchy and sequencing. Even though the task hierarchy structure is not absolute in that a user may deviate from the typical roadmap through the tasks, this hierarchy can be used as a valuable knowledge source to track the discourse focus, manage the displays, and anticipate the needs of the user.

The CUBRICON entity rating module includes a task-dependent representation of the relative inportance of all the entity types known to the system and an algorithm for modifying these ratings depending on task and discourse context. CUBRICON uses a numerical value (on a scale from 0.0 to 1.0) to represent an entity's importance. A pre-defined task-dependent initial assignment of ratings to entities is used when a new task is started by the user. These ratings are modified when the entities are referenced in the dialogue.

3.5 Characteristics Of The Information

The characteristics of the information to be expressed are critical to the selection of an appropriate presentation modality. The following list briefly summarizes CUBRICON's criteria for selecting presentation modality based on characteristics of the information. The CUBRICON design is based on the premise that graphic/pictorial presentation is always desirable.

- o Color-graphics: Selected whenever CUBRICON knows how to represent the information pictorially.
- o Table: Selected when the values of common attribute(s) of several entities must be expressed.
- o Histogram: Selected when a quantitative attribute of several entities must be displayed in a comparative form.
- o Form: A predefined form is selected when the task engaged in by the user requires the form.
- o Printed natural language: Selected for the expression of a proposition, relation, event, or combination thereof that would strain the user's short term memory if speech were used (e.g., technical responses that must be referred to subsequently by the

user); the knowledge structures being expressed are heterogeneous;

- o Spoken natural language: Selected for explanations of displays, verbal highlighting of objects on the displays, informing the user about the system's activity (e.g., "I'm still working" when the user must wait for output from the system), short expressions of relatively non-technical information that can be remembered when presented serially.

3.6 Guidelines For Enhancing Human Understanding

Guidelines from the human factors and psychology disciplines for enhancing human understanding are being incorporated into the CUBRICON system. Smith & Mosier [1986] provide one of the well-known sources of such human factors guidelines. Guidelines that have been incorporated in the CUBRICON system include:

- o Maintain the context of the user/computer dialogue. For example, in managing map displays, CUBRICON tries to retain the most recently discussed or mentioned objects on the displays (avoid premature removal) so as to maintain continuity in the dialogue.
- o Maintain consistency throughout a display. For example, when CUBRICON expands or extends a map display to include some new area, then the same types of objects are displayed in the new region as in the old so that the user does not incorrectly infer that a certain type of object is *not* in the newly added region because it is not being displayed.
- o Maintain consistency across displays. Different displays of the same type should have the same general format. Color keys, titles, help information, etc., should be consistently located for different displays of the same type.

4. MULTI-MEDIA INPUT UNDERSTANDING

People commonly and naturally use coordinated simultaneous natural language (NL) and pointing gestures. These two modes of communication combine synergistically to form an efficient method of expressing definite references and locative adverbials. For example, a person could simply say "this SAM" (surface-to-air missile) and point to an entity on the display to select from among several SAM systems visible on the display. Used in isolation, each of the two modes have shortcomings. If natural language is used alone, then lengthy descriptions are frequently required to identify objects that lack unique "names". For example, to specify a particular SAM from among many SAMs visible on a map display, a person would need say something like "the SAM system at 12.3 degrees longitude and 50.5 degrees latitude" or "the SAM system just east of Kleinburg".

Pointing used alone also has problems: (1) a point gesture can be ambiguous if the point touches the area where two or more graphical figures or icons overlap and (2) the user may inadvertently miss the object at which he intended to point. This latter problem is discussed briefly in the last paragraph of this section. To handle the first problem of ambiguous pointing, some systems use default techniques. These techniques include: (1) a point returns the entity represented by the "top" or "foremost" icon where the system has a data structure it uses to remember the order in which icons are "painted" on the display (i.e., which are further in the background and which are foremost in the foreground); (2) the icons or entities are assigned weights representing importance and the icon with the largest weight is selected as the interpretation of an ambiguous point; or (3) the icon whose "center" is closest to the location pointed at is selected. Combinations of such techniques can also be used. A serious disadvantage of the above listed point-interpretation techniques is that it is difficult, if not impossible, for certain icons to be selected via a point reference.

CUBRICON's acceptance of dual-media input (NL accompanied by coordinated pointing gestures) overcomes the limitations of the above weak default techniques and provides an efficient expressive referencing capability. The CUBRICON methodology for handling dual-media input is a decision-making process that depends on a variety of factors such as the *types* of candidate objects being referenced, their *properties*, the *sentential context*, and the *constraints on the participants* or fillers of the semantic case frame for the verb of any given sentence. CUBRICON's decision-making process draws upon it's knowledge sources discussed in Section 3.

CUBRICON also provides the user with flexibility in several different ways: (1) the user can input natural language via either the speech device or the keyboard, (2) the user is not limited to just one point per NL phrase, but can point several times per phrase, (3) point gestures can occur anywhere within a given NL phrase, and (4) the user can reference four different types of objects via pointing: geometric points, entities represented graphically (e.g., by icons), entries in table displays, and windows on a display.

We present a few brief examples to illustrate the functionality of the system. In each of the following examples, assume that the <point> touches one or more icons.

Example 1: USER: "What is the status of this <point> airbase?"

From the icons touched by the point, the virtual display is searched for the semantic representation of the objects which were graphically displayed by the touched icons. From the hierarchy of the knowledge base, the system determines which of the objects selected by the point gesture are airbases and discards the others.

Example 2: USER: "What is the mobility of this <point> ?"

Example 2 entails the use of the property "mobility" as the critical item of information that is used to determine the referent of the dual-media phrase. After searching the virtual display for the objects touched by the point gesture, CUBRICON determines which of these objects have property "mobility" using the knowledge base of application information.

Example 3: USER: "Remove this window <point>."

In example 3, assuming the <point> touches an icon on a window of the display, the term "window" used with the point gesture insures that the system correctly interprets the reference, which would otherwise be ambiguous.

CUBRICON also includes the ability to handle two types of "ill-formed" input: dual-media expressions which (1) are inconsistent or (2) have an apparent null referent. A dual-media expression is inconsistent when the natural language part of the expression and the accompanying point cannot be interpreted as referring to the same object(s) (e.g., the user says "this airbase" and points to a factory). A dual-media expression has no apparent referent when the user's point touches no icons (i.e. he points to an "empty" area). In both of these cases, CUBRICON infers the intended referent. The CUBRICON methodology for processing multi-media input is discussed in greater detail in [Neal et al., 1988].

5. MULTI-MEDIA OUTPUT GENERATION

The CUBRICON system is being developed so that it embodies the following key features which are essential to maximizing human understanding of presented information: (1) Output presentations should be sensitive to context and relevance. (2) Selection of appropriate presentation media/modalities should be based on the characterisitics of the information to be expressed, alternatives being selected when necessary. (3) The media/modalities should be used in a highly integrated manner during output presentation. (4) Respected human factors guidelines should be adhered to.

The CUBRICON output planning process is highly dependent on the knowledge sources discussed in Section 3. The top level output planning process is summarized below.

1. For each information item or cluster, determine the modality in which it should ideally be expressed. Graphic/pictorial presentation is always desirable. Natural language can always be used as a last resort.
2. Determine whether the resources are available to express the information as desired. Resources: (1) Color graphics display: Are the items to be expressed graphically already on the display? If so, no additions are necessary. If not, is there room to insert them in their "natural" position? (2) Monochrome display: Similar to the color graphics display. (3) Natural language text window: Always available due to scrolling capability. (4) Speech output device: Always available.
3. If the desired display cannot accommodate the new information without modification, determine whether the state of the display can be modified. This depends on the level of importance of the information to be expressed and the task model.

   Possible modifications to the graphics display:

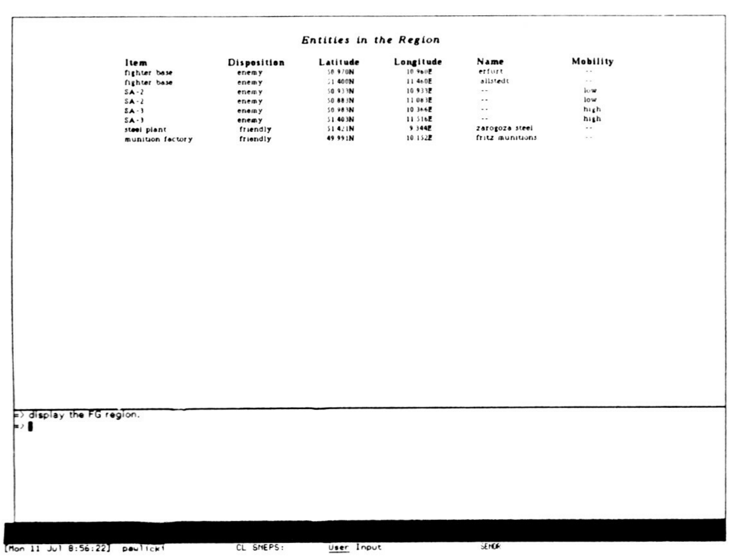
   - o Extend the displayed region to include information items to be expressed.
   - o "Zoom in" to display a portion of the current display in greater detail.
   - o "Pan" to a different region.
   - o Combination of the above.
   - o Open a window on the display to show the new information.
4. If the display status cannot be adequately modified as per step 3 above, try modifying the information to be expressed: trim the amount of information by filtering on the basis of relevance with regard to the user/task model and/or the discourse model.
5. If the information can still not be expressed in the given modality due to insufficient resources for the selected modality, then select another modality and go back to step 2.
6. Finish composing the output having resolved resource restraints.

The following are working examples from a dialogue with CUBRICON to illustrate the functionality of the system. These examples illustrate the output composition process and use of the knowledge sources discussed in the previous sections. The dialogues are concerned with mission planning and situation assessment in a tactical air control domain.

USER: "Display the Inner Fulda Gap Region."

CUBRICON:
Color Graphics Display:

- o Map of Inner Fulda Gap Region with main roads, major cities, waterways, and national boundries (Figure 1).
- o Icons representing entities within the Inner Fulda Gap Region that are above a preset threshold in importance superimposed on the map (Figure 1).



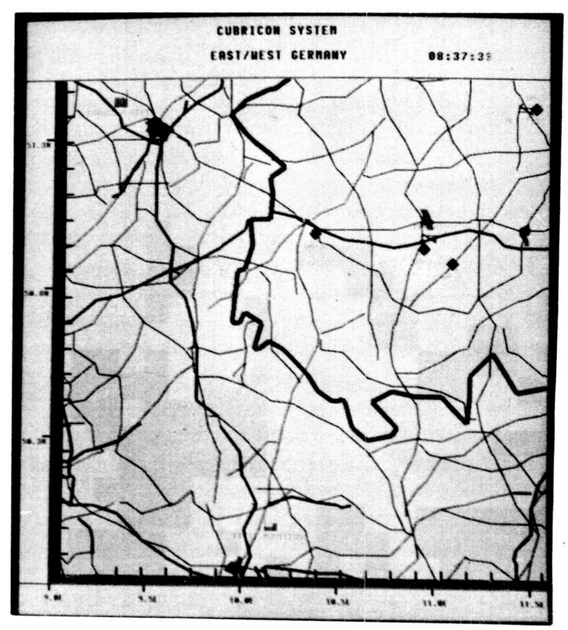


Figure 1. Multi-Modal Presentation Composed by CUBRICON

ORIGINAL PAGE IS
OF POOR QUALITY

Monochrome Graphics Display:

- o Table of relevant attributes for the entities that are displayed on the map display.

DISCUSSION:

CUBRICON's first choice for displaying information is graphically, when appropriate. The Fulda Gap Region boundary is defined in the CUBRICON knowledge base and provides sufficient information for the modality selection process. CUBRICON selects color-graphics as a display modality and generates a map display. Since this was the initial input, the color graphics display was available.

An important feature of CUBRICON is that output presentations are formulated in a context sensitive manner. Entities to be displayed on the map are selected on the basis of their importance to the task at hand. Only entities above a pre-set threshold on the entity rating scale are displayed. Non-domain-specific items such as national borders, roads, rivers, and cities are displayed using the MAP Display System [Hilton, 1987].

Another key feature of the CUBRICON interface system is that it is multi-media. In the present example, the CUBRICON output uses both color graphics (i.e., the map display), and a tabular presentation on the monochrome display. The tabular presentation was selected because there are important attributes associated with the entities displayed on the map display, which would severely clutter the map if included there.

USER: "Where is the Dresden airbase?"

CUBRICON:

Color Graphics Display:

- o Map of Inner Fulda Gap Region with added area that includes the Dresden airbase (Figure 2).
- o Main roads, major cities, waterways, and national boundries (as before but across the whole map, old and new areas) (Figure 2).
- o Icons representing entities within the map area displayed that are above a pre-set threshold in importance are superimposed on the map (same as before but across the whole map, old and new areas) (Figure 2).
- o An airbase icon labelled "Dresden Airbase". This airbase icon is displayed as initially blinking and then highlighted.

Monochrome Graphics Display:

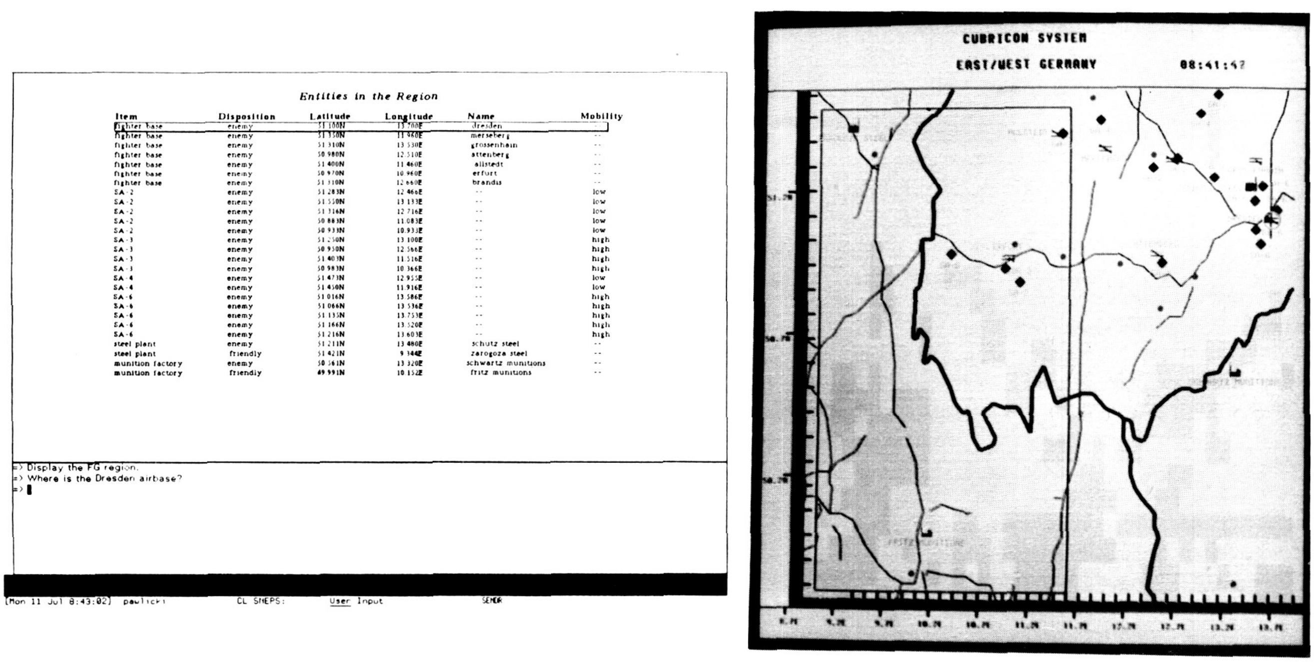
- o Table of relevant entity attributes. Same table as before, but expanded to include new entities added to the map (Figure 2). The Dresden Airbase table entry (row) is highlighted.

DISCUSSION:

The information requested is location. The Dresden airbase has an icon representation and a longitude, latitude associated with it in the CUBRICON knowledge base. The preferred modality for presentation is therefore graphical.

In composing a new map on which to display the Dresden airbase, the system has some choices. These include: open a window on the color-graphics display showing the area around the Dresden airbase, replace the old map on the CRT with a new area around the airbase, or compose a new map including both the old map and the region around the Dresden airbase.

An important principle that CUBRICON tries to follow is to preserve the context of the human-computer dialogue. Since the user task has not changed and there is already a map displayed on the color graphics display, the system expands the displayed area to include the Dresden airbase. CUBRICON selects this option because it provides the requested information while it preserves the display context. When the Dresden airbase is added to the map, the system also blinks it as the system's way of "pointing" to the object under discussion.


Entities in the Region

Item	Disposition	Latitude	Longitude	Name	Mobility
fighter base	enemy	51.100N	13.700E	dresden	--
fighter base	enemy	51.350N	11.960E	merseberg	--
fighter base	enemy	51.310N	13.530E	grossenhain	--
fighter base	enemy	50.980N	12.510E	attenberg	--
fighter base	enemy	51.400N	11.460E	allstedt	--
fighter base	enemy	50.970N	10.960E	erfurt	--
fighter base	enemy	51.310N	12.660E	brandis	--
SA-2	enemy	51.283N	12.466E	--	low
SA-2	enemy	51.550N	13.133E	--	low
SA-2	enemy	51.316N	12.716E	--	low
SA-2	enemy	50.883N	11.083E	--	low
SA-2	enemy	50.933N	10.933E	--	low
SA-3	enemy	51.250N	13.100E	--	high
SA-3	enemy	50.950N	12.566E	--	high
SA-3	enemy	51.403N	11.516E	--	high
SA-3	enemy	50.983N	10.366E	--	high
SA-4	enemy	51.473N	12.955E	--	low
SA-4	enemy	51.450N	11.916E	--	low
SA-6	enemy	51.016N	13.586E	--	high
SA-6	enemy	51.066N	13.536E	--	high
SA-6	enemy	51.135N	13.753E	--	high
SA-6	enemy	51.166N	13.520E	--	high
SA-6	enemy	51.216N	13.603E	--	high
steel plant	enemy	51.211N	13.480E	schutz steel	--
steel plant	friendly	51.421N	9.344E	zarogoza steel	--
munition factory	enemy	50.561N	13.320E	schwartz munitions	--
munition factory	friendly	49.991N	10.152E	fritz munitions	--

Figure 2. Map and Table Maintaining Context and Consistency

Another important guideline to which the CUBRICON system adheres is to maintain consistency throughout a display so as to prevent the user from making false inferences about what is or is not located within the region. In the case of our map display, this means that there should be consistency in the types of objects shown across the entire map. If SAMs are displayed in the old region, then they should be displayed in the newly added map area. Similarly for other types of objects. If this is not done, then the user might incorrectly infer that there were no SAMs in the new area. Guided by the consistency principle, the system also modifies the tabular presentation that is on the monochrome display. The new displays are shown in Figure 2.

USER: "What is the mobility of this SAM (accompanied by a point gesture to a SAM with the mouse device)?"

CUBRICON:
Spoken and Written Natural Language:

- o "The mobility of this SA-3 is high." The phrase "this SA-3" is accompanied by blinking the particular icon as the system's means of pointing to it.

DISCUSSION:

CUBRICON selects natural language as the modality as per the criteria presented in Section 3.5. This example also demonstrates CUBRICON's ability to generate multi-media output that is highly integrated. Specifically, in this example, the spoken natural language definite reference "this SA-3" is coordinated with a blinking reference on the graphics display.

USER: "What are the mobilities of these <point1> <point2><point3><point4><point5>?" The five pointing gestures indicate five different SAM icons on the color map display.

CUBRICON:

- o Color-Graphics Display: The referenced SAMs are highlighted.
- o Monochrome Display: A window containing a table displaying the mobilities of the indicated SAMs is added.

DISCUSSION:

The user's question is almost the same in these last two examples. CUBRICON responds in coordinated NL and pointing when the response is a single proposition, as in the previous example. However, CUBRICON selects a table modality with associated highlighting on the color-graphics display for this last response. This selection is made since the values of a common attribute (mobility) of several entities must be expressed. This situation matches the selection criteria for the table modality as presented in Section 3.5.

6. CURRENT STATUS AND FUTURE DIRECTION

The work discussed in this paper has been implemented on the hardware suite described in Section 2. The knowledge sources discussed in Section 3 have been implemented and the functionality described in this paper has been realized in the current CUBRICON prototype. The examples presented in Sections 4 and 5 are working examples.

The CUBRICON team is continuing its research and development of the concepts and methodology essential to intelligent multi-media human-computer interface systems. This includes continued research and development of the knowledge sources such as the user/task model and discourse model, the automated process of determining the appropriate media/modalities for any given information items or clusters, appropriate modification of the displays including placement of information when needed, the multi-modal composition process, and the role of speech output in a multi-modal interface system.

7. SUMMARY

Modern information processing and decision-aiding systems are complex and require a full range of communication media to facilitate interaction and provide the increased bandwidth for information transfer with the human user. The human-computer interface must be able to use and manage the media and modalities in an intelligent manner. The Intelligent Multi-Media Interface Project is devoted to the development of interface technology that integrates speech, natural language text, graphics, and pointing gestures for human-computer dialogues. The objective of the project is to develop interface technology that uses the media/modalities intelligently in a flexible, context-sensitive, and highly integrated manner modelled after the manner in which humans converse in simultaneous coordinated multiple modalities. As part of the project, a knowledge-based interface system, called CUBRICON is being developed as a research prototype. Several knowledge sources are essential to intelligent use of multiple modalities and are included in our CUBRICON system: a knowledge base of application-specific and related interface knowledge, a definition of the multi-media language in the form of a lexicon and grammar, a discourse model, a user/task model, knowledge of information characteristics and the appropriate corresponding modality for expressing the information, and human factors guidelines for enhancing human understanding and reducing cognitive workload. CUBRICON accepts dual-media input consisting of natural language and simultaneous coordinated pointing gestures. The CUBRICON methodology handles the synergistic mutual disambiguation of simultaneous natural language and pointing as well as inconsistent NL/pointing expressions and expressions that have an apparent null referent. CUBRICON's output composition process includes selection of appropriate modalites and media, determination of whether resources are available, subsequent modification of resources or modification of the information to be expressed (if necessary), modification of selected output media/modalities (if necessary), and composition of the output. Examples were presented to illustrate some of the key functionality of the CUBRICON system.

8. ACKNOWLEDGEMENTS

The other members of the CUBRC team are: D. Funke, S. Glanowski, M. Summers, and S.C. Shapiro.

The MAP Display System was developed by Lt. Mike Hilton (RADC), who graciously allowed us to use it.

9. REFERENCES

1. Arens, Y., Miller, L., Sondheimer, N.K., "Presentation Planning Using an Integrated Knowledge Base," *Proceedings of the Workshop on Architectures for Intelligent Interfaces: Elements and Prototypes*, Lockheed AI Center, Monterey, CA, 1988, pp. 93-107.

2. Carberry, S., "First International Workshop on User Modeling," *AI Magazine*, Vol.8, No.3, 1987, pp. 71-74.

3. Grosz, B. J., "Discourse Analysis," in *Understanding Spoken Language*, D. Walker (ed.), Elsevier North-Holland, New York, 1978, pp. 229-345.

4. Grosz, B.J. & Sidner, C.L., "Discourse Structure and the Proper Treatment of Interruptions," *Proc. of IJCAI*, 1985, pp. 832-839.

5. Grosz, B.J., "The Representation and Use of Focus in a System for Understanding Dialogs," in *Readings in Natural Language Processing*, B.J. Grosz, K.S. Jones, B.L. Webber (eds.), Morgan Kaufmann Pulishers, 1986, pp. 353-362.

6. Hilton, M.L., *Design and Implementation of the MAP Display System*, RADC Report, 1987.

7. Kobsa, A., Allgayer, J., Reddig, C., Reithinger, N., Schmauks, D., Harbusch, K., Wahlster, W., "Combining Deictic Gestures and Natural Language for Referent Identification," *Proceedings of the 11th International Conference on Computational Linguistics*, Bonn, FR Germany, 1986.

8. Neal, J.G., Dobes, Z., Bettinger, K.E., & Byoun, J.S., "Multi-Media References in Human-Computer Dialogue," *Proceedings of the National Conference of the American Association for Artificial Intelligence*, ST. Paul, MN, 1988, (forthcoming).

9. Neal, J.G. & Shapiro, S.C., "Intelligent Multi-Media Interface Technology," *Proceedings of the Workshop on Architectures for Intelligent Interfaces: Elements and Prototypes*, Lockheed AI Center, Monterey, CA, 1988, pp. 69-91.

10. Neches, R. & Kaczmarek, T., *AAAI-86 Workshop on Intelligence in Interfaces*, USC/Information Sciences Institute, August, 1986.

11. Shapiro, S.C., "The SNePS Semantic Network Processing System". In *Associative Networks - The Representation and Use of Knowledge by Computers*, N. Findler (ed.), Academic Press, New York, 1979, pp. 179-203.

12. Shapiro, S.C., "Generalized Augmented Transition Network Grammars for Generation from Semantic Networks," *AJCL*, Vol. 8, No. 1, 1982, pp. 12-25.

13. Shapiro, S.C. & Rapaport, W., "SNePS Considered as a Fully Intensional Propositional Semantic Network," *Proceedings of AAAI-86*, 1986, pp. 278-283; in *Knowledge Representation*, G. McCalla & N. Cercone (eds.), Springer-Verlag Pub.

14. Sidner, C.L., "Focusing in the Comprehension of Definite Anaphora," in *Computational Models of Discourse*, M. Brady & R.C. Berwick (eds.), The MIT Press, 1983, pp. 267-330.

15. Smith, S.L. & Mosier, J.N., *Guidelines for Designing User Interface Software*, MITRE Corporation Technical Report No. 10090, ESD-TR-86-278, 1986.

16. Sullivan, J.W. & Tyler, S.W. (eds.), *Proceedings of the Workshop on Archictectures for Intelligent Interfaces: Elements and Prototypes*, Lockheed AI Center, Monterey, CA, 1988.